第5章

表計算ソフトウエア:可視化(グラフ化)

2022年3月30日

学習目標

- (1) 各グラフの意味を理解する
- (2) 各グラフの作成方法を学習する

1 グラフ化

作成された表を視覚化,グラフ化し、表の形では見えにくい性質をみます.

2 系列

表計算ソフトウェアでは、図1のように、1つのデータの集まりを系列とかデータ系列と呼びます。図1のようなデータ系列になっているとき、各系列の値を項目軸ラベルの値(各都県の値)を比較します。

	項目軸ラベル(表側)			系列名(表頭)]		
	Α /		В		С		D		Е	
1				人	数(単	<u> </u>	1000人	()		
2	都!	果名 (小学生	ſ	中学生	:	高校生	:	大学生	
3(茨	成	18	39	1	10	10	6	37	
4	栃	木	12	28		74	7	7	21	
5	群	馬	12	24		69	6	7	19	
6	埼	玉	4()8	2	17	20	6	114	
7	干	葉	33	39	1:	90	18	3	114	
8	東	京	56	62	3	29	36	5	660	
9	神	奈川	46	31	2	42	23	2	190	
10	合	†	22	11	12	31	123	6	1155	
	小学生の系列				中等	学生	の系列			

図1 系列(列が系列になっている場合)

図2のように行を系列にすることもできます.この場合,小学生数,中学生数,高校生数,大学生数を比較します.

	系列名	В	С	D	E :	項目軸ラベル
1		人	数(単位:	1000人)		
2	都県名	小学生	中学生	高校生	大学生	
3	茨城	189	110	106	37	
4	栃木(128	74	77	21	
5	群馬	124	69	67	19	
6	埼玉	408	217	206	114	栃木の系列
7	千葉	339	190	183	114	
8	東京	562	329	365	660	
9	神奈川	461	242	232	190	東京の系列
10	合計	2211	1231	1236	1155	

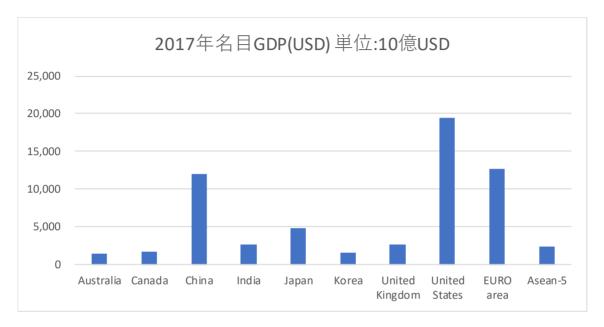
図2 系列を行にした場合

系列について、行と列を入れ替えることができます.

3 棒グラフ

3.1 単純棒グラフ(単一系列の棒グラフ)

棒グラフは、棒の長さが値の大きさを表し、比較するグラフです。図3は、2000年の各国のGDPを表したものです。



動画:Excel 棒グラフの作成

図3 棒グラフの作成手順

- (1) グラフ化する範囲を指定します (上の例: A2:A11 と R2:R11). 単一の系列の場合,系列名を範囲に含めません. また,複数領域の範囲指定なので,2つめの領域は(Ctrl)キーを押しながらマウスで操作します.
- (2) リボンの 挿入 のタブをクリック
- (3) グラフの中から、 縦棒 を選びます.
- (4) 集合縦棒 (最上段左のアイコン)を選びます.
- (5) グラフが表示されるので、タイトルを書きます.

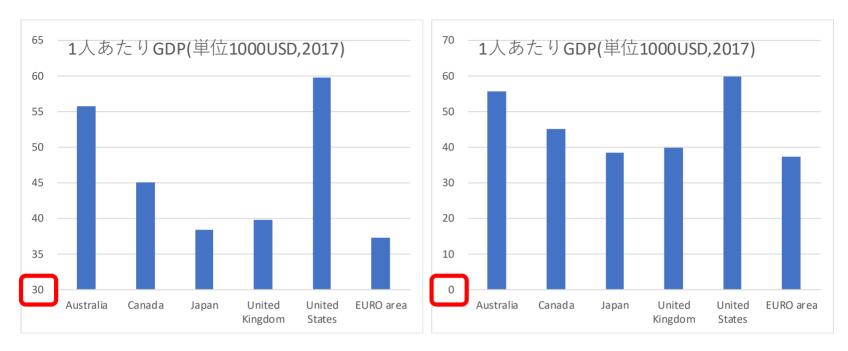


図4 棒グラフ目盛りの調整(左:適切ではない)

図 4 は,縦軸の目盛りを変えたとき,グラフからどのよなイメージを受けるのかの比較です. Australia, Canada, Japan, United Kingdom, United States, EURO area の 7 つの国・地域の 2017 年の 1 人あたりの GDP(1000USD 単位)を比較するためのものです. 左のグラフは,縦軸の最小値を 30 に設定しています. このグラフでは,米国の GDP の棒の長さが日本の約 3 倍になっており,日本が米国の $\frac{1}{3}$ 以下という誤解を与え,適切とはいえません.

そこで、右のように0から始めるようにします。金額のように何倍かに意味のあるデータを比較するときは、できるだけ 目盛りの最小値は0にするようにします。

動画:Excel 目盛りの調整の影響

3.2 集合縦棒

S7.xslx の小学生,中学生,高校生,大学生という4つの系列について,各都県のデータを比較します。主に児童生徒数の合計を比較し、その内訳も比較できます。図5は、集合縦棒グラフで、各系列の色を変えて比較しています。

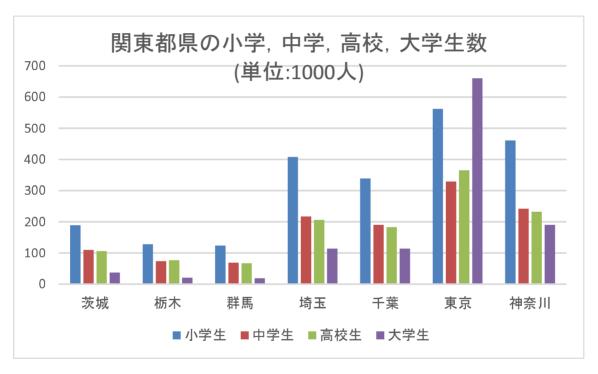


図 5 積み上げ縦棒(単位:1000人)

作成するときは、棒グラフの左上を使います.大学生の数がどの都県が多いのかなどがわかります.動画:Excel 集合縦棒

3.3 積み上げ縦棒

図 6 は、小学生、中学生、高校生、大学生という 4 つの系列について、各都県の人数を積み上げて棒グラフにしたものです。

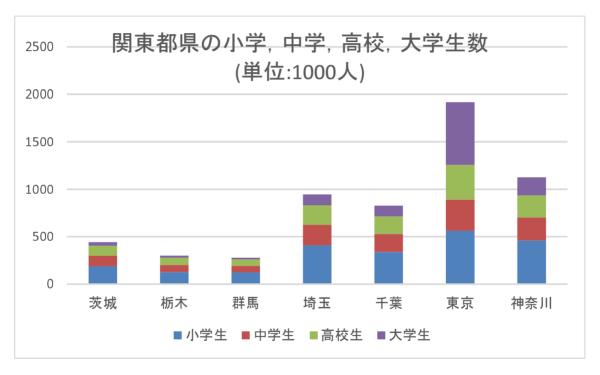


図 6 積み上げ縦棒(単位:1000人)

3.4 積み上げ縦棒

図 7 は、各都県を系列にし、小学生、中学生、高校生、大学生の人数を積み上げて棒グラフにしたものです。 系列の行と列を切り替え:

動画:Excel 棒グラフ 5(積み上げ縦棒, 行列の切り替え)

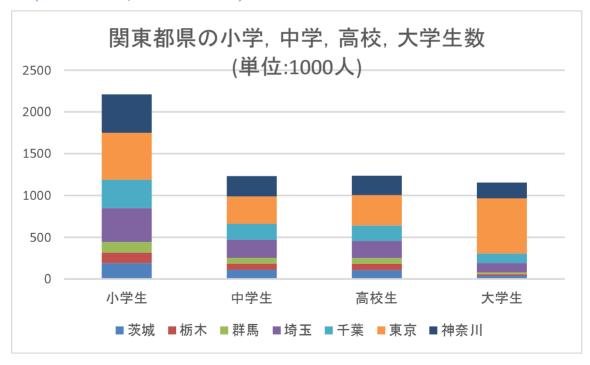


図7 積み上げ縦棒(単位:1000人)行と列を切り替えたもの

3.5 平均線の追加と地図グラフ

■平均線の追加 図 8 は、COVID19 のデータ (S10.xlsx の「全期間累積」) の都道府県毎の感染者数累計を単棒グラフで表現したものです。各都道府県を全国平均 (全国の感染者数÷日本の人口 (百万人単位)) と比較するため、全国の高さに赤い線を追加しました。これにより、全国平均の差異を検討できます。このグラフにより、100 万人あたりの感染者数が多いのは都市部と米軍基地の影響がある都道府県であることがわかります。

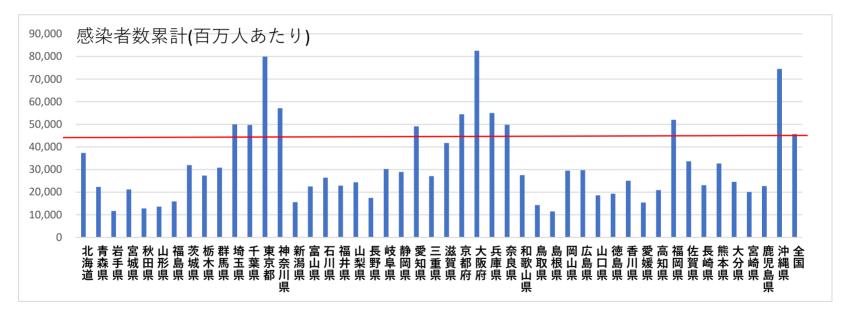


図8 平均線を追加した棒グラフ

平均線は、グラフの描画機能ではなく、図形の機能で作成しました。そのため、作成後のグラフのコピーでは、グラフの部分と線の部分を別々にコピーする必要があります。

- (1) メニューの 挿入 → 図形 → 線
- (2) ドラックアンドドロップで線を描画
- (3) 線を選択し右クリック → 図形の書式設定
- (4) 色や線幅を変更

動画:Excel 平均線を追加

練習のため、図23のような都道府県別の人口100万にあたりの死者数のグラフを作成しましょう.

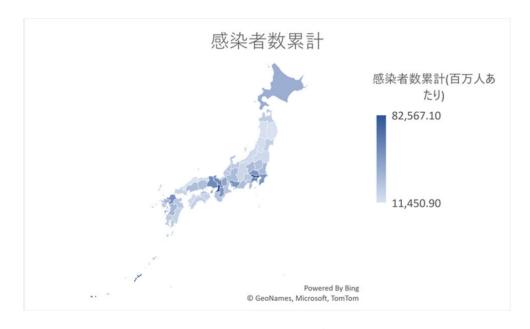


図9 感染者数の地図グラフ

- ■地図グラフ 国別や都道府県別や,市町村別などは,地図上に表示すると位置関係を把握でき,地域別の分析が容易になります。図9は,図8の棒グラフを感染者数が多い都道府県を濃い色で塗り分けた地図グラフです。色の濃さは,周辺の都道府県の色の濃さに影響をうけたり,大きさの倍数を直感的に把握できないというデメリットもあります。Excel 上では,値を知りたい都道府県にマウスカーソルを合わせることで値を参照できます。
 - (1) グラフ化する範囲指定 (図 9 の例では, A6:AV7, 全国は含めません)
 - (2) メニューの 挿入 \rightarrow $\boxed{ マップ } \rightarrow \boxed{ 塗り分けマップ }$

動画:Excel 地図グラフ

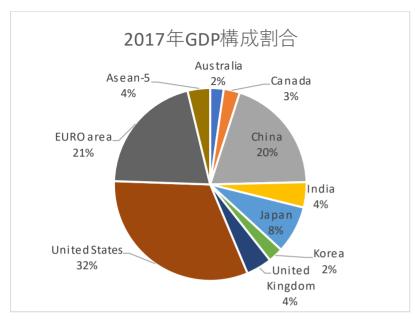
4 割合を比較するグラフ

構成割合の比較するとき, 円グラフや帯グラフが使われます.

4.1 円グラフ (3D 円グラフ)

円グラフは、円弧の面積で、構成割合を比較するものです。図 10 は、比較のため同じデータをグラフ化したものです。 3D 円グラフは、表示面積が見方によって変わるため誤解を与えやすいグラフになります。 たとえば、イギリス (United Kingdom, 4.30%) は、インド(India, 4.26%)に比べて構成割合がかなり大きいように見えます。

棒グラフと同様に、単一の系列の場合系列名を含めません。 クイックレイアウト という機能を使うと簡単に書式設定ができます。動画:Excel 円グラフ



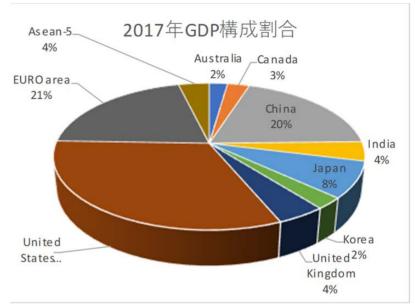


図 10 2017 年の構成割合の比較 (左:円グラフ, 右:3D 円グラフ)

4.2 帯グラフ

複数の構成割合を比較するときは、帯グラフが用いられます(ドーナッツ円グラフが使われることがありますが、内側と 外側の系列では、同じ割合でも半径の違いにより面積が異なるので、誤解を与える可能性があります).

図 11 は、S1.xlsx のシート「GDP share」を使って、グラフ化したものです. 作成するときの注意

• 複数の領域を範囲指定しています. 2 つ目以降は、(ctrl)キーを押しながら範囲指定します.

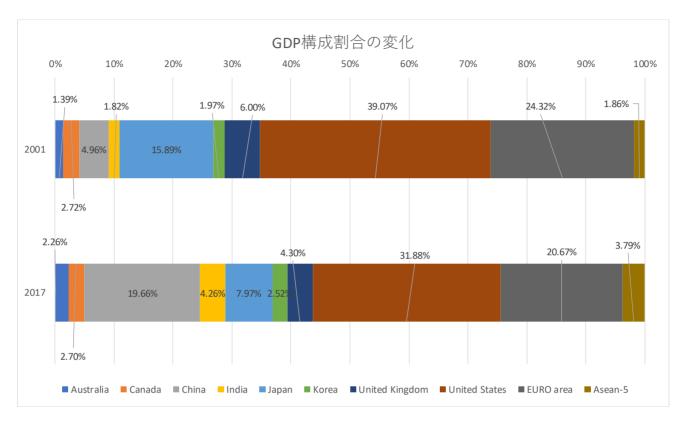


図 11 2001 年と 2017 年 GDP の構成割合の比較

- 系列(年)は、数値になっています。はじめから指定するとその部分も含めてグラフ化されるため、系列は、あとで指定します。
- 2001 年を上に表示させるため、「軸の反転」を使っています.

- 数値が一部重なるため、手作業で文字を移動しています.
- 動画:Excel 帯グラフ(作成)
- 動画:Excel 帯グラフ(軸の逆転)
- 動画:Excel 帯グラフ(データラベル)

5 データの変化を見る(時系列データ)

時間(年,日,秒など)とともに変化するデータを時系列データと言います。時系列データをグラフ化するとき,折れ線グラフを使います。Excel などの表計算ソフトウエアでグラフ化するとき,時を表すデータは,数値や時間,日付などの大きさを表すものにし,「散布図」または「折れ線」でグラフ化します。

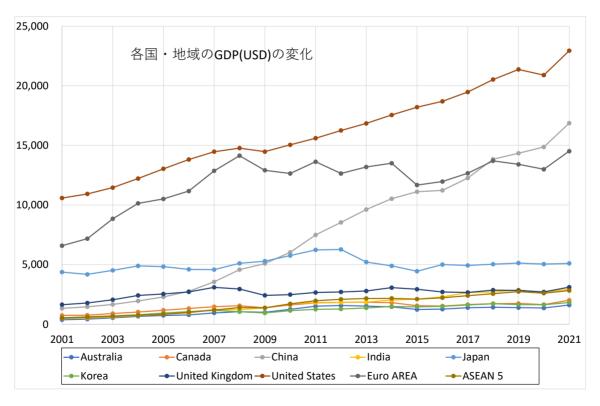


図 12 時系列データのグラフ (散布図で作成した折れ線グラフ)

5.1 散布図で折れ線グラフを作成

次のようにして、図 12 のようなグラフを作成しました.

- 年(数値), 国名を含めて範囲指定(合計や順位の列などは含めない)
- 散布図(直線とマーカー)を利用
- •「グラフの移動」で「新しいシート」にグラフを移動
- 文字のフォントを拡大
- 横軸の範囲を 2001~2021 に変更
- 動画:Excel 折れ線グラフ (散布図で作成)
- 動画:Excel 折れ線グラフ (新しいシートに移動し、グラフの整形)

5.2 一部の系列を抜き出して表示

図 13 では、上位 3 つの国・地域 (China, United States, EURO area) の GDP が大きく、他の国・地域の値の変化が読みにくくなっています。そこで、グラフフィルターを使って、表示する系列(国・地域)を選択します(図 13).

動画:Excel 折れ線グラフ (表示する系列を選択)

グラフのエリアをクリックし選択します.右上に縦に 🛨 🗾 のアイコンが現れるので, 🔟 をクリックします.

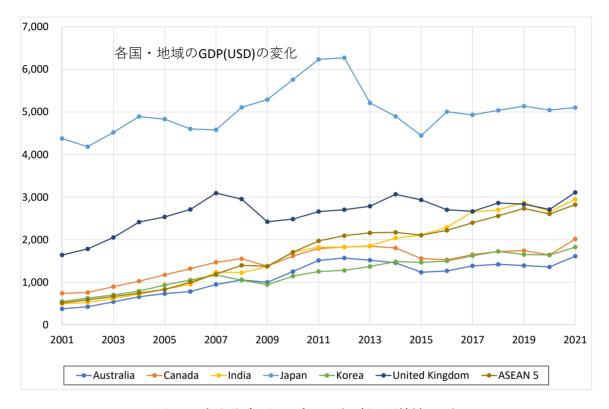


図 13 時系列データのグラフ (一部の国地域のみ)

5.3 時間の間隔

図 14 (出典 総務省統計局「日本の統計 2014」) は、年次毎に日本の総人口を表示したものです(JPN_population.xlsx として、教科書のページからダウンロードできます。

元のデータは、和暦で表示され、平成12年までは、5年間隔、平成15年以降は1年間隔で値が表示されています。

年次(和曆)	総人口(単位1000人)	年次(西暦)	総人口(単位1000人)
昭和20年	72,147	1945	72,147
昭和25年	84,115	1950	84,115
昭和30年	90,077	1955	90,077
昭和35年	94,302	1960	94,302
昭和40年	99,209	1965	99,209
昭和45年	104,665	1970	104,665
昭和50年	111,940	1975	111,940
昭和55年	117,060	1980	117,060
昭和60年	121,049	1985	121,049
平成2年	123,611	1990	123,611
平成7年	125,570	1995	125,570
平成12年	126,926	2000	126,926
平成15年	127,694	2003	127,694
平成16年	127,787	2004	127,787
平成17年	127,768	2005	127,768
平成18年	127,901	2006	127,901
平成19年	128,033	2007	128,033
平成20年	128,084	2008	128,084
平成21年	128,032	2009	128,032
平成22年	128,057	2010	128,057
平成23年	127,799	2011	127,799
平成24年	127,515	2012	127,515

図14 日本の人口

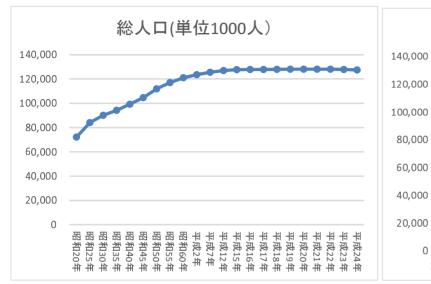




図 15 折れ線グラフ (左) と散布図 (右)

図 15 左は、図 14 左の和暦表示の表から、折れ線グラフ機能を使って描いたものです。マーカーの間隔が昭和の頃は 5 年、 平成 15 年から 1 年間隔になっていますが、意図したものとは異なってしまっています。

図 15 右は、図 14 右の西暦表示の表から、散布図機能を使って描いたものです。マーカーの間隔が昭和の頃は 5 年、平成 15 年から 1 年間隔になってますが、その分、マーカーとマーカーの間隔が $\frac{1}{5}$ になり、目盛りの間隔が調整されています。

2つのグラフを比較すると、和暦表示のグラフの方が長く人口が停滞しているように見えますが、これは、目盛りの間隔が正しく調整されていないことによります。したがって、図 15 左は、誤解を与えるグラフとなります。

5.4 比率尺度, 間隔尺度

棒グラフと同様に、人口を比較するとき、縦軸の原点を注意しなくてはなりません。図 16 は、原点を 50,000 に変更した グラフです。1945 年に比べて 2000 年頃は 3~4 倍になっているという誤解を与えます。ただし、特に小さな差異を見たいと きには、原点を変更することがあります。



図 16 原点を変更したグラフ

- 一般に、その数値が持っている尺度により扱いが異なります.
- 比率尺度 (比例尺度) 金額や発生件数など, 倍数に意味がある尺度です. 100 円は 50 円の 2 倍で, 3000 件は 1000 件の 3 倍で. この場合, 縦軸は 0 から始めます.
- **間隔尺度** 年などなど間隔に意味がある尺度です. 西暦 2000 年と 2005 年の差の 5 年と, 2005 年と 2010 年の差の 5 年の 意味は同じです. 平成 6 年と 15 年では, 名目 GDP を見る上で, 1.5 倍年数が経過しているわけではありません. また, フリースタイルスキーの得点も間隔尺度です. この場合, 縦軸の最小値を調整できます.

5.5 单純移動平均 · COVID-19

 $S10^{**}$.xlsx を使い,全国,各都道府県の COVID-19 の感染者数,死者数日々の変化を見てみたいと思います.このデータは,日付形式での日付は数値として扱えるものです.数値は,1900/1/1 を 1 ととする毎日 1 増える数値で表現されています(日付の表示形式を数値に変更すればその数値を見ることができます).本節では,Excel の「折れ線」を使ってグラフ化します.

	А	В	С	D	Е	F	
1		原系列	原系列				
2	全国	感染者数	死者数				
3	2020/1/16	1	0				
4	2020/1/17	0	0				
5	2020/1/18	0	0				
6	2020/1/19	0	0				
7	2020/1/20	0	0		単純移動平均	単純移動平均	
8	2020/1/21	0	0	全国	感染者数	死者数	
9	2020/1/22	0	0	2020/1/22	0.142857143	0	
10	2020/1/23	0	0	2020/1/23	0	0	
11	2020/1/24	1	0	2020/1/24	0.142857143	0)
12	2020/1/25	1	0	2020/1/25	0.285714286	0)
12	2020/1/26	1	n	2020/1/26	N 128571120	Ω	í

図 17 全国感染者数・死者数推移計算シート

- ■原系列の計算表の作成 図 17 の A,B,C 列に全国もしくはある都道府県のデータを表示させます. このあと,移動平均という加工をします. その加工前の系列のことは「原系列」と呼ばれています. 図 17 のように新しいシート「推移分析」を作成し、グラフを作成していきます.
 - (1) 新しいシート「推移分析」を作成します.
 - (2) 日付を表示させます: A3: =感染者数!A3 A3 をクリックし,「=」を入力し, ワークシート「感染者数」を選択し, A3 をクリック
 - (3) B 列に全国の感染者数を表示させます: B3: =感染者数!AW3 感染者数!AW3 は,全国の 2020/1/16 の感染者数です. 他の都道府県を分析する場合,その都道府県の感染者数の列(例:東京都の場合,感染者数!N3)を指定します.
 - (4) C列に全国の死者数を表示させます: C3: =死者数!AW3 他の都道府県の場ではその都道府県の列を指定
 - (5) A,B,C 列の計算式を最終日付 (2022/3/13) の 790 行まで複写します. 最終日付が異なる場合, その行まで複写します. 複写元: A2:C2 複写先: A3:C790
- ■原系列のグラフの作成 COVID-19 のグラフでは、EXCEL の散布図ではなく EXCEL の折れ線グラフで作成しています. 感染者数と死者数では、大きく(100~1000 倍)異なるので、感染者数と死者数を異なる y 軸を使います.
 - (1) 範囲指定: A2:C790
 - (2) メニューの 折れ線/面グラフの挿入 \rightarrow 2D-折れ線 \rightarrow 折れ線 \mid 折れ線/面グラフの挿入 の左上のグラフ)
 - (3) 感染者数と死者数を異なる軸にします.
 - (4) グラフを選択し、メニューの グラフのデザイン \rightarrow グラフの種類の変更

(5) 図 18 のように,

すべてのグラフ から 組み合わせ を選択

グラフの種類は、感染者数、死者数ともに 折れ線

死者数を第2軸(右の軸)にチェックを入れる

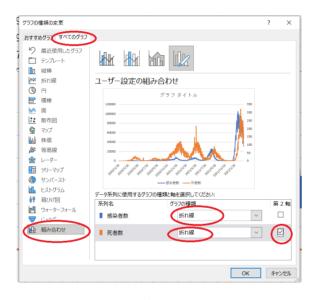


図18 第2軸の設定

■単純移動平均 図 19 をみると細かい周期で振動しているように見えます。これは、曜日により Covid-19 の検査回数が異なりそのため曜日により感染者数が異なることが想定できます。また、死者数も報告日が曜日により異なることが考えられるます。このような曜日のよる周期的な変化を除去するために、1 週間 (7 日間) 周期で変化することを利用し、各日の感染

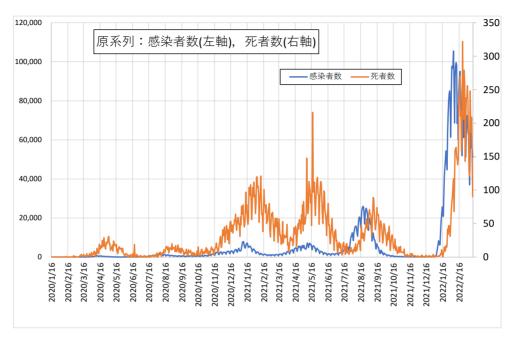


図 19 全国感染者数·死者数推移

者数や死者数をその日を含む前7日間の平均をその日の値とすることが行われます.

このように前?日の平均値で分析することはよく行われ、?日単純移動平均と呼ばれています.このような分析は、株価の推移や外国為替相場の推移などによく用いられます.

■単純移動平均の計算表 図 17 の D~F 列に感染者数, 死者数の 7 日単純移動平均を求めます.

(1)
$$D9: = A9$$

- (2) E9: =AVERAGE(B3:B9) 前7日の平均なので、B3:B9(7個のセル) の平均値
- (3) F9: =AVERAGE(C3:C9) 複写元: D9:F9 複写先: D10:F790 F790 は、いつまでのデータなのかにより異なります.
- ■単純移動平均のグラフ 図 20 のように 7 日単純移動平均をグラフ化してみましょう.
 - (1) <mark>範囲指定: D8:F790</mark> F790 は、いつまでのデータなのかにより異なります.
 - (2) 原系列のグラフの作成と同様にグラフを作成します.
 - (3) 死者数の線を点線にし、色を赤にしました.

図 20 の Y 軸の目盛りは、自動設定のため、図 19 の Y 軸の最大値とは異なります.

- ■Y 軸の目盛りを対数にする(単純移動平均のグラフ) 感染症の感染者数や死者数は、(前週などの) 何倍になるのかで議論されています。このように一定期間(例えば1週間)に倍数で変化するデータは、普通の目盛(線型間隔の目盛)よりも対数目盛を使います。ここでも対数の目盛に変更します。
 - (1) 左の軸を右クリック \rightarrow | 軸の書式設定 $|\rightarrow|$ 軸のオプション $|\rightarrow|$ 対数目盛を表示する にチェックを入れる
 - (2)「負の数値またはゼロは対数グラフに正しくプロットされません」と表示されますが、これは、感染者数に0があるためです。ここでは、無視するのでOKをクリックします(0は、対数では負の無限大になりますのでグラフには表示できません)。
 - (3) 同様に,右の軸を右クリック \rightarrow $\boxed{ 軸の書式設定 } \rightarrow \boxed{ 軸のオプション } \rightarrow \boxed{ 対数目盛を表示する }$ にチェックを入れ、 $\boxed{ OK }$ をクリック

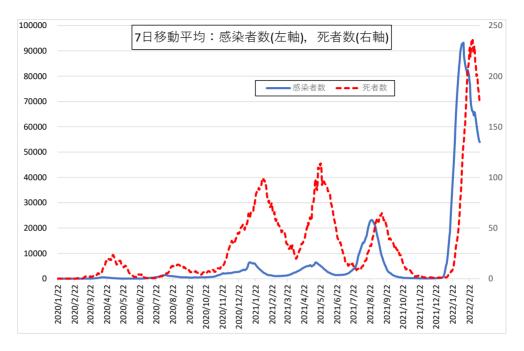


図20 全国感染者数:死者数推移(7日単純移動平均)

動画:Excel 2 軸グラフ・対数目盛

図 21 のように感染者数・死者数の変化がはっきり分かるようになったと思います.次のことをグラフから読み取ってみましょう.

・感染者数が 2022/2/1 頃に頂点をつけているのが「第 6 派」と呼ばれています。感染者数の谷から谷までの期間で,第 1 波から第 6 波の期間を読み取りましょう。グラフにマウスを当てると,日付,その数値が表示されます。図 21 のように,2022/3/13 現在と同様にまだ第 6 波は収束していないかもしれません。その場合,第 6 派の期間の終了は読み



図 21 全国感染者数·死者数推移 (7 日単純移動平均·対数目盛)

取れません.

• 感染者数のピークに比べてと死者数のピークが遅れて出ています. おおよそ何日後か読み取ってみましょう (数値での分析は情報入門 2 で行う予定です).

2つの系列のデータにどのような傾向があるのかを分析するのに、散布図 (XY グラフ) が使われます.

6.1 家計調査の例 (魚介類支出金額と肉類支出金額)

図 22 は、家計調査での魚介類支出金額と肉類支出金額(2020年、二人以上の世帯での年間支出金額)の県庁所在地・政令指定都市別の平均値をプロットしたものです(出典:総務省家計調査、家計調査 家計収支編二人以上の世帯 年報 年次 2020年 の 6. 食料「魚介類」および食料「肉類~乳卵類」から作成)

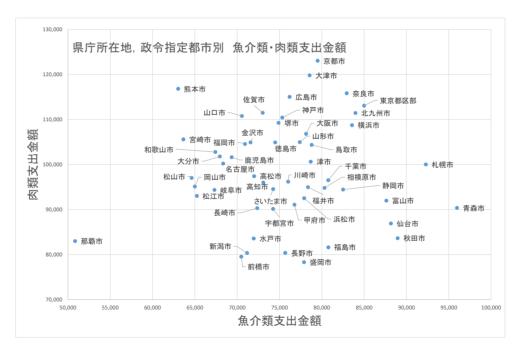


図 22 散布図の例

右方向は、魚介類支出金が多く、上方向は肉類支出金が多いことを表します。したがって、右下の都市は、肉類に比べて 魚介類の支出金額が多く、左上は、魚介類に比べて肉類が多いことを表しています。また、右上は、両支出が多いことを表 してます。これは、支出金額全体が多いことや物価水準が高いことが考えられます、逆に左下はその逆が考えられます。グ ラフを見ると、右下に東北、左上に九州、四国の都道府県が多いことがわかります。

このグラフは,次のようにして,作成しました.

- (1) この表は、教科書のページからダウンロードできます (family_inc_exp_2020.xlsx). をダウンロード
- (2) 散布図の作成. (動画:Excel 散布図の作成3)
- (3) 軸ラベルの設定(動画:Excel 散布図の作成4)
- (4) 軸ラベルの設定(動画:Excel 散布図の作成 5)
- (5) データラベルの設定(動画:Excel 散布図の作成 6) ラベルの位置の修正は、手作業

6.2 COVID-19 の 100 万人あたり死者数の要因分析(感染者数と死亡率の散布図)

図 23 は、都道府県別 100 万人あたりの死者数者数です。

一部の都道府県が全国平均赤線の約2倍以上と大きくまっています。この原因について、2つの要因に分けて分析をしたいと思います。

感染者死亡率 = $\frac{100 \, \text{万人あたりの死者数}}{100 \, \text{万人あたりの感染者数}}$

100 万人あたりの死者数 = 100 万人あたりの感染者数×感染者死亡率

となり、100万にあたりの死者数が多いのは、100万にあたりの感染者数が多いのか、感染者死亡率が多いのかまたが その 両方であるのか分析できます。そこで、各都道府県の人口 100万にあたりの感染者数と感染者死亡率の散布図を作成します。

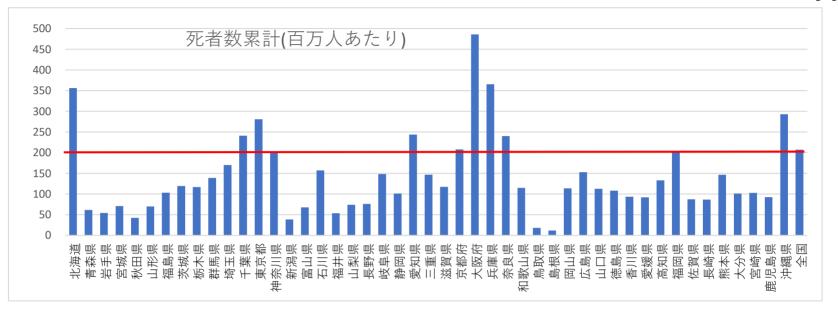


図 23 都道府県別 100 万人あたりの死亡者数

- (1) \$10 の 全期間累積 のシート
- (2) A7:AW7 を範囲指定(感染者数)
- (3) キーボードから(Ctrl)キーを押しながら A9:AW9 を範囲指定(感染者死亡率)
- (4) 散布図の散布図でグラフを作成
- (5) データラベルの追加で、都道府県名を追加する(文字が重なるので適当に移動いさせる)
- (6) 全国の値で、線を追加し、4つのエリアに分ける.

動画:Excel ポジショニング分析

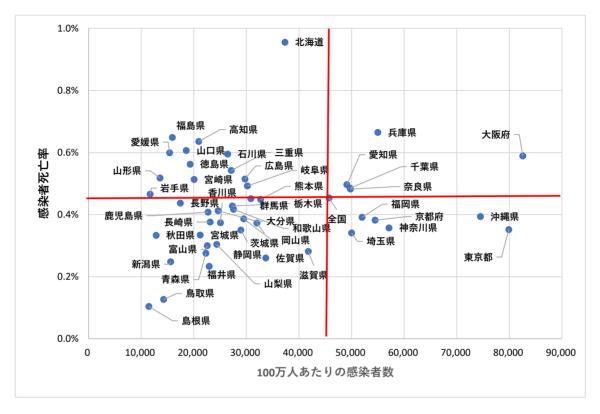


図 24 人口 100 万にあたりの感染者数と感染者死亡率

図 24 は、右上方向に人口 100 万にあたりの死者数が多い都道府県になります。第 1 象限は、人口 100 万にあたりの感染者数と感染者死亡率がともに高いことが要因になっており、第 2 象限は、感染者死亡率が高いこと第 4 象限は、人口 100 万にあたりの感染者数が多いことが要因になっています。このように、サンプル(都道府県)を 2 次元の点で示し、サンプルがどのような位置にあるのか分析することはよく行われます(ポジショニング分析)。

7 練習問題 (グラフ)

7.1 練習問題 S1-9

- (1) 2001 年から 2017 年にかけての 2001 年を 100 とする GDP の指数を比較する折れ線グラフ (散布図) を作成しなさい.
- (2) (1) の折れ線グラフ (散布図) から、中国、インド、Asean-5 の系列を除いた折れ線グラフ (散布図) を作成しなさい.
- (3) 2001 年(横軸) と 2017 年(縦軸) の一人あたりの GDP を比較する XY グラフ (散布図) を作成しなさい.

7.2 練習問題 S2-5

- (1) 2013年の4社の売上高を比較する棒グラフを作成しなさい.
- (2) 2009 年から 2013 年の売上高の変化を比較する折れ線グラフ(散布図)を作成しなさい.

7.3 練習問題 S4-3

ターン点,エア点,スピード点の合計を比較する積み上げ棒グラフを作成しなさい(図 25).

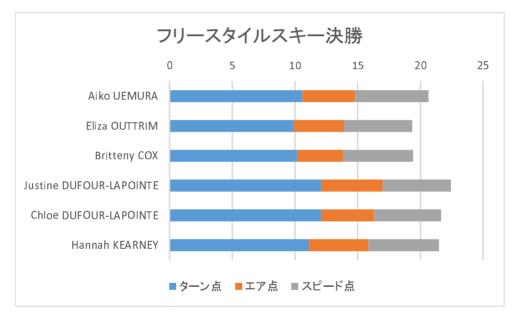


図 25 積み上げ横棒グラフ

7.4 練習問題 S5-2

(1) 日本, アメリカ, イギリス, ドイツ, フランスの 2010 年の人口構成を比較するグラフ (図 26)

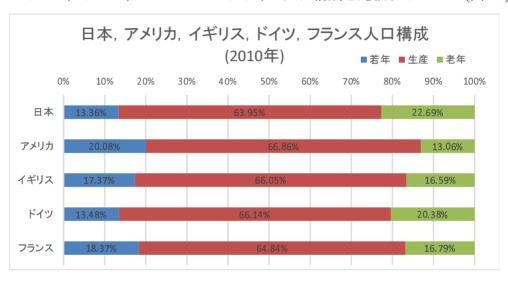


図 26 人口構成

(2) 図 27 のような生産年齢人口の増加率を比較するグラフを作成しなさい.

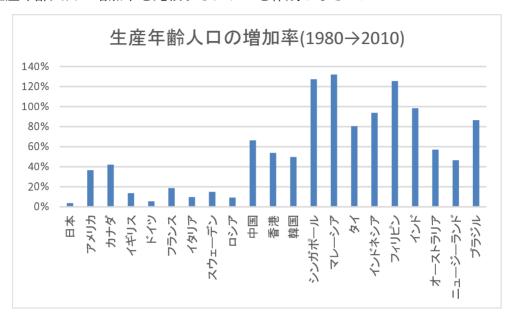


図 27 生産年齢人口の増加率